

RAW SEQUENCE LISTING

DATE: 02/06/2001

PATENT APPLICATION: US/09/265,585A

TIME: 12:27:43

Input Set : A:\Pto.amc

Output Set: N:\CRF3\02062001\I265585A.raw

P.5

4 <110> APPLICANT: Philip N. BENFEY
 5 Laura Di LAURENZIO
 6 Joanna WYSOCKA-DILLER
 7 Jocelyn E. MALAMY
 8 Leonard PYSH
 9 Yrjo HELARIUTTA
 10 Jun LIM
 12 <120> TITLE OF INVENTION: Scarecrow Gene, Promoter and Uses Thereof
 14 <130> FILE REFERENCE: 5914-066-999
 16 <140> CURRENT APPLICATION NUMBER: 09/265,585A
 C--> 17 <141> CURRENT FILING DATE: 2001-01-18
 19 <150> PRIOR APPLICATION NUMBER: 08/842,445
 20 <151> PRIOR FILING DATE: 1997-04-24
 22 <150> PRIOR APPLICATION NUMBER: 08/638,617
 23 <151> PRIOR FILING DATE: 1996-04-26
 25 <160> NUMBER OF SEQ ID NOS: 144
 27 <170> SOFTWARE: FastSEQ for Windows Version 4.0
 29 <210> SEQ ID NO: 1
 30 <211> LENGTH: 2163
 31 <212> TYPE: DNA
 32 <213> ORGANISM: Arabidopsis thaliana
 34 <400> SEQUENCE: 1
 36 ccttatttat aaccatgcaa tctcagcacc aacaaccctt caatctccat ggcggaatcc 60
 38 ggcgatttca acggtggtca acctctctct catagtctct tgagaacaac ttcttccggt 120
 40 agtagcagca gcaacaaccg tggctcctct cctcctctct ctccctcttt agtgatggtg 180
 42 agaaaaagat tagcttccga gatgtcttct aaccctgact acaacaactc ctctcgctct 240
 44 cctcgccgtg tctctcacct tcttgactcc aactacaata ctgtcacacc acaacaacca 300
 46 ccgtctctta cggcgccggc tactgtatct tctcaacca acccaccact ctctgtttgt 360
 48 ggcttctctg gtcttcccggt ttttcttcca gaccgtggtg gtcggaatgt tatgatgtcc 420
 50 gtacaacca tggatcaaga ctcttcatct tcttctgctt cacctactgt atgggttgac 480
 52 gccattatca gagaccttat ccattcctca acttcagtct ctattcctca acttatccaa 540
 54 aacgttagag acattatctt cccttgtaac ccaaattctg gtgctcttct tgaatacagg 600
 56 ctccgatctc tcatgtctct tgatccttcc tcttctctg acccttctcc tcaaactttc 660
 58 gaacctctct atcagatctc caacaatctt tctcctccac aacagcaaca gcagcaccaa 720
 60 caacaacaac aacagcataa gcctcctct cctccgattc agcagcaaga aagagaaaat 780
 62 tcttctaccg atgcaccacc gcaaccagag acagtgcagg ccactgttcc cgcggtccaa 840
 64 acaaatacgg cggaggcttt aagagagagg aaggaagaga ttaagaggca gaagcaagac 900
 66 gaagaaggat tacaccttct cacattgctg ctacagtgtg ctgaagctgt ctctgctgat 960
 68 aatctcgaag aagcaaacaa gcttcttctt gagatctctc agttatcaac tccttacggy 1020
 70 acctcagcgc agagagtagc tgcttacttc tcggaagcta tgtcagcgag attactcaac 1080
 72 tegtgtctcg gaatttaacg ggctttgcct tcacggtgga tgcctcaaac gcatagcttg 1140
 74 aaaatggtct ctgcggttca ggtctttaat gggataagcc ctttagtgaa attctcacac 1200
 76 ttacagcga atcaggcgat tcaagaagca tttgagaaag aagacagtgt acacatcatt 1260
 78 gacttggaac tcatgcaggg acttcaatgg cctggtttat tccacattct tgcttctaga 1320
 80 cctggaggac ctccacacgt gcgactcacg ggacttgga cttccatgga agctottcag 1380
 82 gctacaggga aacgtctttc ggatttcaca gataagcttg gcctgccttt tgagttctgc 1440
 84 ccttttagctg agaaagttgg aaacttgga actgagagac tcaatgtgag gaaaagggaa 1500

RAW SEQUENCE LISTING

DATE: 02/06/2001

PATENT APPLICATION: US/09/265,585A

TIME: 12:27:43

Input Set : A:\Pto.amc

Output Set: N:\CRF3\02062001\I265585A.raw

```

86 gctgtggctg ttcactggct tcaacattct ctttatgatg tcaactggctc tgatgcacac 1560
88 actctctggg tactccaaag gtaaaataaa cattaccttt taatcaactct ttatctataa 1620
90 attattttta gattatatag gaaagatatg ttctaaaaag ctggcttttt tgggtaaatga 1680
92 ttggggaatg aacagattag ctctctaaagt tgtgacagta gtggagcaag atttgagcca 1740
94 cgctgggttct ttcttaggaa gattttaga ggcaatacat tactactctg cactctttga 1800
96 ctcaactggg gcaagctacg gcgaagagag tgaagagaga catgtcgtgg aacagcagct 1860
98 attatcgaaa gagatacgga atgtattagc ggttgaggga ccacgcagaa gcggtgaagt 1920
100 gaagtttgag agctggaggg agaaaatgca acaatgtggg tttaaaggta tatctttagc 1980
102 tggaaatgca gctacacaag cgactctact gttgggaatg ttctcttcgg atgggttacac 2040
104 tttggttgat gataatggta cacttaagct tggatggaaa gatctttcgt tactcaactgc 2100
106 ttcagcttgg acgcctcgtt cttagttttc ttctcttttt tcacaaacaa tgtgcccata 2160
108 aat 2163
111 <210> SEQ ID NO: 2
112 <211> LENGTH: 653
113 <212> TYPE: PRT
114 <213> ORGANISM: Arabidopsis thaliana
116 <400> SEQUENCE: 2
118 Met Ala Glu Ser Gly Asp Phe Asn Gly Gly Gln Pro Pro Pro His Ser
119 1 5 10 15
121 Pro Leu Arg Thr Thr Ser Ser Gly Ser Ser Ser Asn Asn Arg Gly
122 20 25 30
124 Pro Pro Pro Pro Pro Pro Pro Leu Val Met Val Arg Lys Arg Leu
125 35 40 45
127 Ala Ser Glu Met Ser Ser Asn Pro Asp Tyr Asn Asn Ser Ser Arg Pro
128 50 55 60
130 Pro Arg Arg Val Ser His Leu Leu Asp Ser Asn Tyr Asn Thr Val Thr
131 65 70 75 80
133 Pro Gln Gln Pro Pro Ser Leu Thr Ala Ala Thr Val Ser Ser Gln
134 85 90 95
136 Pro Asn Pro Pro Leu Ser Val Cys Gly Phe Ser Gly Leu Pro Val Phe
137 100 105 110
139 Pro Ser Asp Arg Gly Gly Arg Asn Val Met Met Ser Val Gln Pro Met
140 115 120 125
142 Asp Gln Asp Ser Ser Ser Ser Ala Ser Pro Thr Val Trp Val Asp
143 130 135 140
145 Ala Ile Ile Arg Asp Leu Ile His Ser Ser Thr Ser Val Ser Ile Pro
146 145 150 155 160
148 Gln Leu Ile Gln Asn Val Arg Asp Ile Ile Phe Pro Cys Asn Pro Asn
149 165 170 175
151 Leu Gly Ala Leu Leu Glu Tyr Arg Leu Arg Ser Leu Met Leu Leu Asp
152 180 185 190
154 Pro Ser Ser Ser Ser Asp Pro Ser Pro Gln Thr Phe Glu Pro Leu Tyr
155 195 200 205
157 Gln Ile Ser Asn Asn Pro Ser Pro Pro Gln Gln Gln Gln Gln His Gln
158 210 215 220
160 Gln Gln Gln Gln Gln His Lys Pro Pro Pro Pro Ile Gln Gln Gln
161 225 230 235 240
163 Glu Arg Glu Asn Ser Ser Thr Asp Ala Pro Pro Gln Pro Glu Thr Val
164 245 250 255

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/265,585A

DATE: 02/06/2001

TIME: 12:27:43

Input Set : A:\Pto.amc

Output Set: N:\CRF3\02062001\I265585A.raw

```

166 Thr Ala Thr Val Pro Ala Val Gln Thr Asn Thr Ala Glu Ala Leu Arg
167      260      265      270
170 Glu Arg Lys Glu Glu Ile Lys Arg Gln Lys Gln Asp Glu Glu Gly Leu
171      275      280      285
173 His Leu Leu Thr Leu Leu Leu Gln Cys Ala Glu Ala Val Ser Ala Asp
174      290      295      300
176 Asn Leu Glu Glu Ala Asn Lys Leu Leu Leu Glu Ile Ser Gln Leu Ser
177 305      310      315      320
179 Thr Pro Tyr Gly Thr Ser Ala Gln Arg Val Ala Ala Tyr Phe Ser Glu
180      325      330      335
182 Ala Met Ser Ala Arg Leu Leu Asn Ser Cys Leu Gly Ile Tyr Ala Ala
183      340      345      350
185 Leu Pro Ser Arg Trp Met Pro Gln Thr His Ser Leu Lys Met Val Ser
186      355      360      365
188 Ala Phe Glu Val Phe Asn Gly Ile Ser Pro Leu Val Lys Phe Ser His
189      370      375      380
191 Phe Thr Ala Asn Gln Ala Ile Gln Glu Ala Phe Glu Lys Glu Asp Ser
192 385      390      395      400
194 Val His Ile Ile Asp Leu Asp Ile Met Gln Gly Leu Gln Trp Pro Gly
195      405      410      415
197 Leu Phe His Ile Leu Ala Ser Arg Pro Gly Gly Pro Pro His Val Arg
198      420      425      430
200 Leu Thr Gly Leu Gly Thr Ser Met Glu Ala Leu Gln Ala Thr Gly Lys
201      435      440      445
203 Arg Leu Ser Asp Phe Thr Asp Lys Leu Gly Leu Pro Phe Glu Phe Cys
204      450      455      460
206 Pro Leu Ala Glu Lys Val Gly Asn Leu Asp Thr Glu Arg Leu Asn Val
207 465      470      475      480
209 Arg Lys Arg Glu Ala Val Ala Val His Trp Leu Gln His Ser Leu Tyr
210      485      490      495
212 Asp Val Thr Gly Ser Asp Ala His Thr Leu Trp Leu Leu Gln Arg Leu
213      500      505      510
215 Ala Pro Lys Val Val Thr Val Val Glu Gln Asp Leu Ser His Ala Gly
216      515      520      525
218 Ser Phe Leu Gly Arg Phe Val Glu Ala Ile His Tyr Tyr Ser Ala Leu
219      530      535      540
221 Phe Asp Ser Leu Gly Ala Ser Tyr Gly Glu Glu Ser Glu Glu Arg His
222 545      550      555      560
224 Val Val Glu Gln Gln Leu Leu Ser Lys Glu Ile Arg Asn Val Leu Ala
225      565      570      575
227 Val Gly Gly Pro Ser Arg Ser Gly Glu Val Lys Phe Glu Ser Trp Arg
228      580      585      590
230 Glu Lys Met Gln Gln Cys Gly Phe Lys Gly Ile Ser Leu Ala Gly Asn
231      595      600      605
233 Ala Ala Thr Gln Ala Thr Leu Leu Gly Met Phe Pro Ser Asp Gly
234      610      615      620
236 Tyr Thr Leu Val Asp Asp Asn Gly Thr Leu Lys Leu Gly Trp Lys Asp
237 625      630      635      640
239 Leu Ser Leu Leu Thr Ala Ser Ala Trp Thr Pro Arg Ser

```

RECEIVED**FEB 08 2001****TECH CENTER 1600/2900**

RAW SEQUENCE LISTING DATE: 02/06/2001
 PATENT APPLICATION: US/09/265,585A TIME: 12:27:43

Input Set : A:\Pto.amc
 Output Set: N:\CRF3\02062001\I265585A.raw

```

240                      645                      650
243 <210> SEQ ID NO: 3
244 <211> LENGTH: 23
245 <212> TYPE: PRT
246 <213> ORGANISM: Arabidopsis thaliana
248 <400> SEQUENCE: 3
250 Pro Ala Val Gln Thr Asn Thr Ala Glu Ala Leu Arg Glu Arg Lys Glu
251 1                      5                      10                      15
253 Glu Ile Lys Arg Gln Lys Gln
254                      20
257 <210> SEQ ID NO: 4
258 <211> LENGTH: 23
259 <212> TYPE: PRT
260 <213> ORGANISM: Saccharomyces cerevisiae
262 <400> SEQUENCE: 4
264 Leu Lys Arg Ala Arg Asn Thr Glu Ala Ala Arg Arg Ser Arg Ala Arg
265 1                      5                      10                      15
267 Lys Leu Gln Arg Met Lys Gln
268                      20
271 <210> SEQ ID NO: 5
272 <211> LENGTH: 23
273 <212> TYPE: PRT
274 <213> ORGANISM: Arabidopsis thaliana
276 <400> SEQUENCE: 5
278 Arg Arg Leu Ala Gln Asn Arg Glu Ala Ala Arg Lys Ser Arg Leu Arg
279 1                      5                      10                      15
281 Lys Lys Ala Tyr Val Gln Gln
282                      20
284 <210> SEQ ID NO: 6
285 <211> LENGTH: 23
286 <212> TYPE: PRT
287 <213> ORGANISM: Mus musculus
289 <400> SEQUENCE: 6
291 Ile Arg Arg Glu Arg Asn Lys Met Ala Ala Ala Lys Cys Arg Asn Arg
292 1                      5                      10                      15
294 Arg Arg Glu Leu Thr Asp Thr
295                      20
297 <210> SEQ ID NO: 7
298 <211> LENGTH: 23
299 <212> TYPE: PRT
300 <213> ORGANISM: Homo sapiens
302 <400> SEQUENCE: 7
304 Arg Lys Arg Met Arg Asn Arg Ile Ala Ala Ser Lys Cys Arg Lys Arg
305 1                      5                      10                      15
307 Lys Leu Glu Arg Ile Ala Arg
308                      20
311 <210> SEQ ID NO: 8
312 <211> LENGTH: 23
313 <212> TYPE: PRT

```

RAW SEQUENCE LISTING

DATE: 02/06/2001

PATENT APPLICATION: US/09/265,585A

TIME: 12:27:43

Input Set : A:\Pto.amc

Output Set: N:\CRF3\02062001\I265585A.raw

```

314 <213> ORGANISM: Homo sapiens
316 <400> SEQUENCE: 8
318 Val Arg Leu Met Lys Asn Arg Glu Ala Ala Arg Glu Cys Arg Arg Lys
319 1 5 10 15
321 Lys Lys Glu Tyr Val Lys Cys
322 20
325 <210> SEQ ID NO: 9
326 <211> LENGTH: 23
327 <212> TYPE: PRT
328 <213> ORGANISM: Zea mays
330 <400> SEQUENCE: 9
332 Lys Arg Lys Glu Ser Asn Arg Glu Ser Ala Arg Arg Ser Arg Tyr Arg
333 1 5 10 15
335 Lys Ala Ala His Leu Lys Glu
336 20
338 <210> SEQ ID NO: 10
339 <211> LENGTH: 23
340 <212> TYPE: PRT
341 <213> ORGANISM: Zea mays
343 <400> SEQUENCE: 10
345 Met Arg Gln Ile Arg Asn Arg Asp Ser Ala Met Lys Ser Arg Glu Arg
346 1 5 10 15
348 Lys Lys Ser Tyr Ile Lys Asp
349 20
352 <210> SEQ ID NO: 11
353 <211> LENGTH: 23
354 <212> TYPE: PRT
355 <213> ORGANISM: Oryza sativa
357 <400> SEQUENCE: 11
359 Arg Arg Met Val Ser Asn Arg Glu Ser Ala Arg Arg Ser Arg Lys Lys
360 1 5 10 15
362 Lys Gln Ala His Leu Ala Asp
363 20
365 <210> SEQ ID NO: 12
366 <211> LENGTH: 43
367 <212> TYPE: PRT
368 <213> ORGANISM: Arabidopsis thaliana
370 <400> SEQUENCE: 12
372 Ala Phe Glu Lys Glu Asp Ser Val His Ile Ile Asp Leu Asp Ile Met
373 1 5 10 15
375 Gln Gly Leu Gln Trp Pro Gly Leu Phe His Ile Leu Ala Ser Arg Pro
376 20 25 30
378 Gly Gly Pro Pro His Val Arg Leu Thr Gly Leu
379 35 40
382 <210> SEQ ID NO: 13
383 <211> LENGTH: 43
384 <212> TYPE: PRT
385 <213> ORGANISM: Arabidopsis thaliana
387 <400> SEQUENCE: 13

```

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/265,585A

DATE: 02/06/2001
TIME: 12:27:44

Input Set : A:\Pto.amc
Output Set: N:\CRF3\02062001\I265585A.raw

L:17 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:443 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:16
L:462 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:17
L:516 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:18
L:1180 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:31
L:1262 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:32
L:1336 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:33
L:1398 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:34
L:1578 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:39
L:1603 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:40
L:1606 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:40
L:1642 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:42
L:1645 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:42
L:1648 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:42
L:1660 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:42
L:1685 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:43
L:1691 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:43
L:1700 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:43
L:1855 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:46
L:1931 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:48
L:2018 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:50
L:2083 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:52
L:2194 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:54
L:2380 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:56
L:2452 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:57
L:2454 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:57
L:2482 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:57
L:2582 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:58
L:2603 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:58
L:2606 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:58
L:2648 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:58
L:2672 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:59
L:2694 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:60
L:2696 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:60
L:2698 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:60
L:2715 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:61
L:2718 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:61
L:2721 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:61
L:2745 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:62
L:2771 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:63
L:2804 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:64
L:2806 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:64
L:2826 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:65
L:2829 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:65
L:2932 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:67
L:2934 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:67
L:2937 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:67
L:2940 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:67

VERIFICATION SUMMARY

DATE: 02/06/2001

PATENT APPLICATION: US/09/265,585A

TIME: 12:27:44

Input Set : A:\Pto.amc

Output Set: N:\CRF3\02062001\I265585A.raw

L:2943 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:67
L:2946 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:67
L:2952 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:67